PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/31, C07K 14/31, C12N 1/21, C07K 16/12, A61K 39/085

A2

(11) International Publication Number: WO 95/34655

(43) International Publication Date: 21 December 1995 (21.12.95)

(21) International Application Number: PCT/US95/07100

(22) International Filing Date: 5 June 1995 (05.06.95)

(30) Priority Data:
08/254,968 7 June 1994 (07.06.94) US

(71) Applicant: THE ROCKEFELLER UNIVERSITY [US/US]; 1230 York Avenue, New York, NY 10021 (US).

(72) Inventors: CHEUNG, Ambrose; Apartment #10C, 500 East 63rd Street, New York, NY 10021 (US). FISCHETTI, Vincent, A.; 440 Jaon Court, West Hempstead, NY 11552 (US).

(74) Agent: REA, Teresa, Stanek; Burns, Doane, Swecker & Mathis, P.O. Box 1404, Alexandria, VA 22313-1404 (US).

(81) Designated States: CA, JP, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: SURFACE PROTEIN OF STAPHYLOCOCCUS AUREUS

(57) Abstract

Methods and compositions are provided for preventing Staphylococcus aureus infections associated with the use of catheters and similar devices. They are based on the neutralization of ability of surface proteins of staphylococcis to adhere to the surface of the catheters at the skin-catheter junction as the first step in the systemic invasion of the host by the organism.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**GB**	United Kingdom	**MR**	Mauritania
AU	Australia	**GE**	Georgia	**MW**	Malawi
BB	Barbados	**GN**	Guinea	**NE**	Niger
BE	Belgium	**GR**	Greece	**NL**	Netherlands
BF	Burkina Faso	**HU**	Hungary	**NO**	Norway
BG	Bulgaria	**IE**	Ireland	**NZ**	New Zealand
BJ	Benin	**IT**	Italy	**PL**	Poland
BR	Brazil	**JP**	Japan	**PT**	Portugal
BY	Belarus	**KE**	Kenya	**RO**	Romania
CA	Canada	**KG**	Kyrgystan	**RU**	Russian Federation
CF	Central African Republic	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CG	Congo			**SE**	Sweden
CH	Switzerland	**KR**	Republic of Korea	**SI**	Slovenia
CI	Côte d'Ivoire	**KZ**	Kazakhstan	**SK**	Slovakia
CM	Cameroon	**LI**	Liechtenstein	**SN**	Senegal
CN	China	**LK**	Sri Lanka	**TD**	Chad
CS	Czechoslovakia	**LU**	Luxembourg	**TG**	Togo
CZ	Czech Republic	**LV**	Latvia	**TJ**	Tajikistan
DE	Germany	**MC**	Monaco	**TT**	Trinidad and Tobago
DK	Denmark	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**US**	United States of America
FI	Finland	**ML**	Mali	**UZ**	Uzbekistan
FR	France	**MN**	Mongolia	**VN**	Viet Nam
GA	Gabon				

SURFACE PROTEIN OF STAPHYLOCOCCUS AUREUS

FIELD OF THE INVENTION

This invention relates to methods and compositions for preventing infections by Staphylococcus aureus (S. aureus), especially those triggered by the use of catheters prosthetic devices and heart valve replacements. More specifically, it relates to methods and compositions useful for inhibiting the ability of a surface protein of S. aureus to implement adhesion of the organism to endothelical cells or to catheters at the skin-catherter junction thereby initiating infection.

BACKGROUND OF THE INVENTION

S. aureus is one of the most frequently encountered pathogens in infections acquired in the hospital. It accounts for 25% of all hospital acquired infections resulting from the use of catheters or similar structures. Since a great majority of patients entering a hospital require some sort of intravenous device, there is a very high probability of infection. The same type of risk applies to the use of prosthetic devices such a as hip and other joint replacements because the ability of staphylococci to adhere to such devices.

One of the initial reactions of the mammalian host to the presences of a catheter is to coat the object with fibrinogen and other matrix proteins as a prelude to the systemic reaction which is intended to expel the invasion. In a hospital setting, this provides an opportunity of infection by S. aureus which attaches

itself to the fibrinogen at the skin-catheter injunction and thereafter works its way through the skin and into the blood. Since the strains prevalent in hospitals, nursing homes and other patient care facilities are often antibiotic resistant, these types of infections are extremely serious and very difficult to contain. Accordingly, the art has expended much effort to prevent such opportunistic infections.

BRIEF SUMMARY OF THE INVENTION

A surface protein, and the gene which expresses it, have now been discovered. This protein enables the invading bacteria to adhere to the fibrinogen. Antibodies to fibrinogen or to the surface protein will prevent bacterial adhesion and thereby inhibit infection. The protein and segments of the protein are useful as vaccines or for the production of antibodies useful for passive protection of patients prior to the use of a catheter or equivalent device.

This invention, therefore, comprises the protein itself and segments thereof, the gene and segments thereof which produce such products, vectors for the gene and its useful segments, organisms transformed by such vectors, monoclonal and polyalonal antibodies to the protein and its useful segments, vaccines produced utilizing the protein and its segments and methods of preventing S. aureus infections utilizing such products. The invention also includes diagnostic probes utilizing the gene products described herein.

THE FIGURES

There follows a brief description of the figures.

Fig. 1 Western blot of cellular fractions of clone number 14 probes with fibrinogen followed by anti-fibrinogen antibody conjugate. The control contains the lysate of an E. coli clone with pBR322 insert in ϒ Zap. The arrows indicate a 34 kD reactive band together with an upper band which may be a dimer. Results of a duplicate blot probed with ^{125}I fibrinogen were similar.

Fig. 2 Western blot (A) and silver-stained gel (B) of the periplasmic extract of clone number 14 fractionated in a fibrinogen column. The Western blot was probed with fibrinogen/anti-fibrinogen conjugate. The short arrow indicates the fibrinogen-reactive band from crude lysate positive control. The long arrow marks a 34 kD protein that reacts with fibrinogen [4th lane in (A)] and is eluded with 3 M potassium thiocyanate, but not with PBS with 0.5 M NaCl nor with acid elusion.

Fig. 3 Is the complete sequence of clone 36 (C36) encoding fibrinogen reactive protein.

Fig. 4 Sequence comparison of strain DB to that of coagulases from strains 8325-4, 213 and BB with the Pileup program under the GCG package. The arrows indicate the 11 amino acid sequence that is unique to the protein. This sequence shares homology with a cell wall anchor motif found in other gram

positive wall protein However, there is no complete identify to this motif in the Genbank. The arrows at resiues 409 and 419 indicate a unique sequence segment of this protein.

The amino half (residues 59-325) of this protein is primarily helical as predicted by the Garnier analysis. Two areas (residues 58-194 and 264-297) reveal a 7 residue periodicity in which residues in position 'a' and 'd' in a heptad motif 'abcdefg' are either hydrophobic or nonpolar. This pattern is consistent with a coiled-coil conformational structure. The sequence is analyzed by the Matcher Program (22).

Fig. 5 Sequence comparison of our protein (strain DB) to that of coagulases from strains 8325-4, 213 and BB with the Pileup program under the GCC package. The arrows indicate the 11 amino acid sequence that is unique to our protein. This sequence shares homology with a cell wall anchor motif found in other gram positive cell wall protein. However, there is no complete identify to this motif in the Genbank.

The following abbreviations are employed in the description of this invention:

Strain DB	- a wild strain of S. aureus
N2Y broth	- a commercially available growth medium
LB	- a commercially available growth medium
IPTG	- Isopropyl-beta-D-thiogalactopyronoside
BSA	- bovine serum albumin

X-gal	- 5-bromo-4-chlero-3-indolyl-beta-D-galactopyranoside
SDS	- sodium dodecyl sulfate
NP40	- a commercially available nonionic detergent
pBR322	- a commercially available plasmid of known structure
pBluescript	- a commercially available phagemid of known structure
SSPE	- 0.15NaCl, 10mM NaH_2PO_4, Im MEDTA pH7.4
TSE buffer	- 0.1M Tris, 20% sucrose, 5mM EDTA pH8
PBS	- phosphate buffered saline

Strain DB has been deposited at the American Type Colture Collection under the accession number _____.

The following Materials and Methods section is provided for convience and ease of understanding of this invention.

MATERIALS AND METHODS

Bacteria, plasmids and phage

Strain DB which has been phenotypically characterized (5) was used in the construction of a genomic library. E. coli strain Sure (Stratagene) was the host cell for the λZap vector. Clones 14 and 36 were phagemids derived from fibrinogen-reactive

plaques containing the insert. A pBluescript phagemid, pAC8, which contained a protein A clone was derived from a Zap genomic library of DB.

Media and Antibiotics

Unless otherwise indicated, the following media wew used: NZY broth and LB (23) were used for the growth of E. coli strains.

Preparations of affinity purified goat anti-fibrinogen conjugate

Goat anti-human fibrinogen antibody obtained commercially (Cappel, West Chester PA) was affinity-purified on a fibrinogen column as previously described (7). The fibrinogen column was prepared with glutardialehyde activated beads (Boehringer Mannheim, Indianapolis, IN) as described in the manufacturer's insert (4). The monospecificity of the affinity purified goat anti-human fibrinogen IgG was verified by an immunoblot using purified fibrinogen and plasma as antigens (7). The protein concentration of the affinity purified antibody was determined by the BCA protein assay reagent (Pierce Chemicals, Rockeford, IL). Affinity purified goat anti-fibrinogen antibody was conjugated to bovine intestinal alkaline phosphates (Sigma, St. Lousis, MO) as described by Voller (32).

Construction and screening of staphylococcal genomic library

A genomic library of strain DB was constructed with the Zap vector that has been digested with EcoRI and dephosphorylated (Stratagene cloning kit). DB chromosomal DNA was extracted from

lysostaphin-lysed cells as previously described (5,29). Genomic DNA was sheared with a 26 gauge syringe and subject to gel filtration on Sepharose CL2B to remove fragments smaller than 1kb. Fractions containing 4-5 kb fragments were pooled, treated with T4 polymerase to produce blunt ends, methylate with EcoRI methylase (New England Biolabs, Boston, MA), insert into the EcoRI site of the Zap vector with EcoRI linkers and packaged in vitro with Gigapack packing extracts (Statagene). Over 90% of the recombinant phages were recognized as white plaques when plated on lac^- host strain Sure in the presence of IPTG and X-gal.

For the screening of fibrinogen reactive plaques, recombinant phage was incubated with the E. coli host on NZY agar (23) at 42°C for gene expression. Following transfer to duplicate nitrocellulose filters (82 mm in diameter, Schleicher & Schuell, Keene, NH), the filters were blocked with 10ml of TNT buffer (10 mM Tris with 0.15 M NaCl and 0.05% Tween 20) containing 1% BSA for 1 hour at RT and then incubated at 37°C for 3h with 25 ug of fibrinogen (Sigma #4883). This fibrinogen preparation had been further purified over a protein A sepharose column to remove contaminated IgG and was found to be essentially free of contaminants as determined by a silver-stained SDS-gel containing this protein. The filters were then washed once with TNT containing 0.1% BSA and 0.1% NP40, and twice with TNT with 0.1% BSA for 5 min. each. Affinity purified goat anti-fibrinogen antibody alkaline phosphatase conjugate diluted 1:1000 in TNT buffer with 1% BSA was then incubated with the filter for 1 hour at RT. After washing the filters twice with TNT with 0.1% BSA and 0.1% NP40 and three times with TNT with 0.1% BSA for 5 min each, fibrinogen-reactive plaques were visualized with 5-bromo-

4-cloro-3-indolyl phosphate as a substrate as described by Blake et al (1). A λZap vector with a pBR322 insert plated on E.coli or E. coli cells alone served as negative controls. Positive clones detected by fibronogen/anti-fibrinogen conjugate were confirmed by allowing plaques on duplicate filters to react with ^{125}I fibrinogen. This screening procedure was similar to the method described above except that ^{125} fibrinogen (100,000 cpm) was used in place of the cold unlabeled fibrinogen. Following reaction, the nitrocellulose filters were washed twice with TNT with 0.1% BSA and 0.1% NP40 and three times with TNT with 0.1% BSA for 5 min each, and finally subjected to autoradiography. Purified plaques were isolated by rescreening each positive clone at least 4 times.

DNA sequencing of the fibrinogen reactive clones

By infecting the E. coli host strain simultaneously with a f1 helper phage (R408) and the λZap vector containing the insert, a single strand DNA containing the pBluescript phagemid with the insert can be packaged for recirculation, thereby generating subclones in E. coli cells when plated in LB/ampicillin agar (Statagene cloning kit instructions). Plasmids for sequencing were purified from E. coli by equilibrium centrifugation in Cesium chloride-ethidium bromide gradients (23). The purity of the plasmid was confirmed by digestion with restriction enzymes (New England BioLab, Beverly, MA). By using both T3 and T7 primers flanking the insert, plasmid sequencing of clones 14 and 36 were performed with the Sequenase kit (U.S. Biochemicals, Cleveland, OH) following the manufacture's instruction (27). Additional primers were obtained for sequencing from within the insert.

Southern blot hybridization

Southern blot hybridization was performed with random primed samples of gel-purified DNA fragments as probes (12,23). Briefly, chromosomal DNA digested with restriction enzymes was resolved on 0.7% TBE gel and transferred onto Hybond-N^+ membrane Amersham, Arlington Heights, IL) (14). DNA probes were labeled with ^{32}P ($-^{32}P$ deoxycytidine triphosphate, Amersham) using the random primed DNA labeling kits (Boehringer Mannheim). The membrane was then hybridized with the ^{32}P labeled DNA probe at 65°C overnight, washed twice with 2X SSPE with 0.1% SDS at RT for 10 min each followed by 1X SSPE with 0.1% SDS at 65°C for 15 min. The membrane was then subject to autoradiography with an intensifying screen at -70°C.

Expression of fibrinogen-reactive protein of S. aureus in E. coli

One of the fibinogen-binding clones, clone 14 was evaluated for the expression of the fibrinogen binding protein in E. coli. E. coli cells containing this clone was grown in 10ml of LB with 50 ug/ml of ampicillin at 37°C until the OD_{600} reached 1.0. Cells were collected by cetrifugation at 7,000 g for 10 min and resuspended in 1.25 ml of ice cold TSE buffer (100 mM Tris, pH 8.0, containing 20% sucrose and 5mM EDTA). Lysozyme was added to a final concentration of 0.5 mg/ml and the sample was iced for 20 min. For whole cell lysate, 0.25 ml of this suspension was removed and 7.5 ul of Triton X-100 and 50 ul of DNase solution (10 mM $MgCl_2$ with 100 ug/ml of DNase) were added. The sample was frozen (-70°C) and thawed twice.

Magnesium chloride was added to the remaining 1 ml. cell suspension (50 mM final concentration) to stabilize the spheroplasts which were then pelted at 7,000 g for 15 min. The supernatant was filtered through a 0.45 um Millipore membrane to obtain the periplasmic fraction.

To lyse the spheroplasts, 0.25 ml of DNase solution was added to the pellet along with 0.75 of water. The spheroplasts were aspirated vigorously several times with a pasteur pipette, frozen and thawed twice as described above. The lysate generated by this treatment was centrifuged for 49,000 g for 1 hr. The supernatant filtered through a 0.4 um membrane was designated cytoplasmic fraction.

The transformed E. coli has been deposited at the American Type Culture Collection under the accession number _____.

SDS-PAGE and immunoblot analysis

Cellular extracts (10 ul each) were separated on 9% SDS-polyacrylamide gel slabs by the method of Laemmli (21). Prestained molecular standards (BRL, Gaithersburg, MD) were run concurrently in adjacent wells. After electrophoresis, the gel was either stained with silver (Pierce Chemicals) or transferred onto nitrocellulose (30). After transfer, the nitrocellulose filters were allowed to react with fibrinogen followed by affinity purified goat anti-fibrinogen alkaline phosphatase conjugate and reactive substrate as described for the screening of the genomic library. In some experiments, the filters were incubated with ^{125}I fibrinogen (250,000 cpm), washed and autoradiographed as described above.

To detect other proteins in the cell extracts, the specific antibody diluted in blocking buffer (Tris 10mN with 0.5 M NaCl and 0.05% Tween 20, pH 8.2) was incubated with the blot for 2 hours at RT. This was followed by incubation with an appropriate alkaline phosphate conjugate for 1 hour and then processed for band visualization as previously described (3).

Partial purification of the fibrinogen-reactive protein of S. aureus

In an attempt to purify the fibrinogen-reactive protein from clone 14, periplasmic extracts from 4 L of culture were prepared as previously described. Briefly, cells were harvested by centrifugation (7,000 g for 20 min) and resuspended in 75 ml of TSE buffer containing 37.5 mg of lysozyme. After incubation on ice for 20 min, $MgCl_2$ was added to a final concentration of 50 mM and spheroplasts segmented at 7,000 g for 30 min. The periplasmic fraction was aspirated from the supernatant, filtered through a 0.45 um membrane and immediately applied to a fibrinogen column (2.5 x 20 cm) followed by rotation at $4^{o}C$ overnight. The fibrinogen column was prepared by mixing 25 mg of fibrinogen and 5 gm of glutardialdehyde beads as described (4). After collecting the fall through, the column was washed with 150 ml. of PBS followed by 150 ml of PBS with 0.5 NaCl. The fibrinogen-binding protein was then eluded by rotating the column with 10 ml of 3 M potassium thiocyanate at RT for 20 Min followed by collection. In preliminary studies, a similar elusion procedure with 0.1M glycine pH 3.0 was not successful. Fractions from the column were concentrated in a Centricon 10 (Amicon, Danvers, MA) and analyzed by SDS-PAGE and immunoblots with fibrinogen as described.

Computer analysis of sequence data

DNA protein sequence analysis, and sequence comparison with database were conducted with the Sequence Analysis Software Package from the Genetics Computer Group (University of Wisconsin, Madison, WI) (9). The deduced amino acid sequence of the putative protein was compared to a sequence database by the algorithm of Pearson and Lipman (TFASTA implementation of GenBank) (25). The fibrinogen reactive protein sequence shown in the figures has been assigned to clone 36.

RESULTS

Isolation of fibrinogen reactive clones

Using the foregoing procedures, a λ Zap library of strain DB, was screened for clones that were reactive with fibrinogen. Of 100,000 plaques screened, three novel clones, designated 14, 30 and 36, were found to be highly reactive with both ^{125}I fibrinogen and fibrinogen/antifibrinogen conjugate on immunoblots. Subclones containing the pBluescript phagemid together with the insert were subsequently generated in E. coli strain Sure. Plasmid DNA from alkaline lysis minipreps of clones 14, 30 and 36, upon digestion with EcoRI which released the inserts, revealed DNA fragments of 4.6, 3.6 and 3.2kb, respectively. Using the 4.6, 3.6 and 3.2 kb fragments as separate probes, Southern blot analysis of Eco RI digests of these clones established that they hybridized each other. These clones did not hybridize with the EcoRI fragments of pAC8, a protein A probe of DB, thus eliminating the possibility of a false positive reaction between expressed protein A gene product

and goat anti-fibrinogen antibody conjugate during the screening procedure. Further analysis indicated that clone 14 comprises about 2/3 of the mature molecule, C 36, extending into the C-terminus.

Expression studies of the fibrinogen reactive protein of S. aureus

Based on restriction analysis, clones 14 and 30 were similar. Although clone 36 contained the complete gene as determined by sequence analysis, expression of the fibrinogen reactive protein with this clone was found to be difficult. Notably, a culture of clone 14 when grown to late stationary phase (i.e. OD_{600mm} 1.5) also resulted in a significantly decreased yield in fibrinogen-reactive protein. This result can be explained either by toxicity of this protein on E. coli or by increased proteolytic breakdown during stationary phase. For these reasons, the expression of the partial protein was evaluated in clone 14 that has been grown to early stationary phase (OD_{600mmm}=1.0).

Expression studies of different fractions from clone 14 with Western Blots probed with either ^{125}I fibrinogen or fibrinogen/anti-fibrinogen conjugate established that the protein, which has a molecular size of 34 kD was found in the whole cell, periplasmic and membrane fraction (Fig 1). In contrast, a crude lysate of an E. coli clone which contained a pBluescript phagemid with a pBR322 insert did not react with fibrinogen (Fig.1). With some fractions (e.g. membrane), there was also a higher molecular weight band, possibly a dimer, which reacted with fibrinogen. Neither of these proteins were fusions

proteins as they were not inducible with IPTG. Additionally, these bands did not react with polyclonal and monoclonal anti-beta-galactosidase antibody (1:1000 dilution) (Boehringer Mannheim) on immunoblots. The fibrinogen reactive band also did not react with affinity purified chicken anti-protein A antibody (Accurate Chemicals, Westbury, NY), thus providing additional evidence that protein A was not cloned which, by binding to IgG, could lead to false positive results.

To confirm the binding specificity of this protein to fibrinogen, periplasmic extracts which contained fewer contaminating bands were harvested from 4 L of E. coli cells expressing the protein of clone 14 and applied to an affinity column with fibrinogen linked beads. The cloned proteins of interest, as analyzed by silver stain and Western blots with ^{125}I fibrinogen and fibrinogen/anti-fibrinogen conjugate, was found in precolumn fractions and the 3M potassium thiocyanate eluant. However, they were not found in the fall-through, PBS with 0.5 M NaCl eluant, nor in the acid eluant (glycine pH 3.0) Fig. 2). Although the protein was not purified to homogeneity in this one step procedure (Fig. 2), these results clearly indicate the binding specificity of this protein to fibrinogen.

Sequence analysis of the fibrinogen reactive protein

The complete sequence of the fibrinogen protein expressed by clone 30 is shown in Fig. 3. The sequence revealed an open reading frame of 1,935 nucleotides. The sequence has a guanosine-cytosine (GC) content of 34.7%, in contrast to the 30% GC content in the staphylococcal genome (10). The higher GC content is attributable to the carboxy terminal half of the

molecule (39.7%). Putative transcription and translation signals and ribosomal binding sites are indicated in Fig. 3. The first 26 amino acids have features characteristic of a bacterial signal peptide (16). Based on the predicted cleavage site, the mature protein has a predicated size of 69,991 Da with a deduced pI of 6.5.

Analysis of the deduced amino acid sequence revealed three distinct domains in this protein. With the exception of residues 27-58, the N-terminal half (residues 58-325) of the protein is primarily helical as predicted by the Garnier analysis (15). Two areas (residues 58-194 and 264-294) within the helical portion of the molecule reveal a 7 residue periodicity in which residue in positions 'a' and d' in a heptad motif 'abcdefg' are either hydrophobic or nonpolar (Fig. 4). This finding is suggestive of a stable coiled-coil conformational structure in these areas (14.22). The second domain between residues 326 and 505 denotes a proline and glycine rich region (20%). Of the 180 residues present, there are 17 proline and 19 glycine residues. This contrasts with the N-terminal portion of the molecule in which only 3% of the residues are either proline or glycine while the remaining carboxyl portion reveals a composition of 14% proline/glycine residues. The carboxyl-terminal domain (residues 506-645) consists of 5 tandem, direct repeats of 27 amino acids each followed by 5 terminal amino acids (PRVTK). Divergence is observed mainly in the outside repeats. Conformational analysis indicated that this repeat region is nonhelical and contains mostly beta-sheets. Comparison of the protein sequence with others in the Genbank database revealed significant homology to three published S. aureus coagulases from S. aureus strains 8325-4, BB and 213 (18,19,26). With the exception of residue 7

in the primary translation product of the fibrinogen reactive protein, the N-terminal 33 amino acid residue which include the leader peptides among all four sequences were identical and therefore are likely to possess identical sequence cleavage sites (see Fig. 3). Comparing residues 1 to 422 in the fibrinogen reactive protein to coagulases in strains 8325-4, BB and 213, there are 56.2%, 73.2% and 56.2% identity, respectively. The identity between residues 423 and 645 comprising the five repeated units to homologous regions in the coagulase sequences increased markedly to 93.9%. 95.1% and 96.9% for strains 8325-4, BB and 213, respectively (Fig. 5). Like that of fibrinogen-reactive protein, the C-termini of coagulase sequences of strains 8325-4, BB and 213 are composed of repeating units of 27 homologous, but no identical, amino acids followed by the terminal sequence PRVTK (Fig. 6). However, the number of repeating units differ among strains. Although the fibrinogen reactive protein sequence displayed features that are common to the coagulase sequence,a careful comparison revealed a unique stretch of 11 amino acids between residues 409 and 419 (SVTLPSITGES) in the middle of the proline/glycine rich region (Fig. 5). Of interest is the fact that the motif LPSITGES shares homology with a cell wall anchor motif (LPXTGX) found in other gram positive surface protein (12,28). However, there is no complete identity to this heptad motif among sequences in the Genbank database.

Based on all of the foregoing, it is clear that a novel fibrinogen reactive protein of S. aureus has been cloned. This protein is both structurally and functionally different from other apparently similar proteins such as the coagulases. The protein and the gene which expresses it are illustrated in Fig.

3. The figure shows the complete clone 36 and the protein it expresses. Clone 14 runs from nucleotide 684 to nucleotide 1935 in Fig. 3. The protein expressed by clone 14 runs from amino acid residue 229 to 645. Clone 30 is substantially the same as clone 14 and expresses substantially the same protein.

Sequence analysis clearly indicates that the fibrinogen-reactive proteins of this invention shares significant homology with staphylococcal coagulases. Recent evidence by Boden and Flock also suggested that the fibrinogen binding protein of S. aureus may possess cross-reactivity with anti-coagulase antibody (2). However, several lines of evidence indicate that clone 36 expresses a unique fibrinogen-binding protein as do clones 14 and 30. First, expression studies of clone 14 which expresses amino acid residues 229 to 645 show that the expressed protein is necessary for fibrinogen binding. In contrast, classical coagulase has been found to complex with prothrombin to from staphylothrombin which subsequently converts fibrinogen to fibrin (10,17). Secondly, there is a unique stretch of 11 amino acids in the sequence (residues 409-419) that are not found in any of coagulase sequence described. Third, this unique amino acid sequence shares homology with a cell wall anchor motif (LPXTGX) that is found to be necessary for anchoring in a variety of gram positive surface proteins (13,28). Based on these findings, it would appear that the fibrinogen reactive protein may belong to a family of coagulase-like proteins, yet it is both structurally and functionally distinct from any of the coagulases previously described.

The isolation of fibrinogen binding protein from staphylococcal whole cell lysates with conventional chromatographic methods has been reported in two studies (11,31). However, the molecular weight (420 vs 62 kD) and the amino acid composition differs widely between the two studies. In contrast to the 62 kD protein, methionine and tyrosine, but not cysteine residues, are present in the protein of this invention. The fibrinogen reactive protein described here as expressed by clone 36 comprises predominately lysine (11.2%), threonine (9.3%) and glutamic acid (9%) while glycine (16.8%), glutamic acid (15%) and lysine (13.3%) were the abundant amino acids in the 62 kD protein (31). In addition, the deduced isoelectric point (pI-6.5) of the fibrinogen reactive protein also differs from the basic pI (about 10.2) of the 62kD protein.

Previous studies have revealed that the fibrinogen binding component of S. aureus is a cell wall constituent because it is absent in staphylococcal L form (10) and because bacterial clumping in the presence of fibrinogen is abolished upon whole cell digestion with proteinases (4). In reviewing the molecular architecture of the C-terminal region of other gram positive surface-anchored proteins, it is evident that they contain several conserved features (13,14). In the C-terminal of these proteins, a charged tail (4-7 amino acids) is usually preceded by a highly hydrophobic membrane anchor (about 16-20 amino acids), the hexamer LPXTGX, a proline-glycine rich domain and a C-terminal repeat region (14). Clearly, the C-terminal region of the fibrinogen reactive protein is different from the model described. In particular, the region preceding the stop codon lacks a charged tail and a hydrophobic membrane anchor. Instead, the five terminal amino acids (PRVTK) are preceded by five

repeats of 27 amino acid each. In addition, the region N-terminal to these repeats is a broad proline/glycine region (residues 326-505) in the middle of which is a unique sequence (LPSITGE) that shares homology with a cell wall anchor motif (LPXTGX) found in other gram positive surface proteins. Notably, this molecular architecture at the C-terminus is similar to those described for pneumococal surface protein A (35). The amino terminal half of pneumococcal surface protein A, like that of fibrinogen reactive protein, is α-helical and is consistent with an α-helical coiled protein conformation. The α-helical region is followed by a proline-rich domain and a repeat domain consisting of ten 20-amino-acid repeats. In addition, it also lacks a classic membrane anchor and a charged tail. In contrast to the fibrinogen reactive protein, however, there is no LPXTGE motif in the C-terminal region of pneumococcal surface protein A.

The novel proteins of this invention are surface proteins of S. aureus. No such proteins have previously been detected, isolated and characterized. They are principally characterized by their ability to bind fibrinogen. The molecular weight of the protein expressed by clone 36 is about 69,991 Da and its isoelective point 6.5. The gene which expresses this protein contains about 1935 nucleotides. Other characteristic features of the protein and of segments C14 and C30 are described above. Because of the difficulty in expressing protein from clone 36, the preferred clone of this invention is clone 36. The preferred fibrinogen binding protein is the protein expressed by this clone.

The protein of this invention, as specifically described herein, will be recognized by the skilled artisan as representative of a class of protein which may differ amongst the various strains of S. aureus, but will all be characterized as having substantially the same number of amino acid residues, the same tertiary structure and binding activity. They may differ slightly in the identity of the amino acid residues at specific positions in the protein chain. All such proteins are included within the scope of this invention.

The genes which generate the proteins of this invention may also differ slightly amongst various strains, but they all have the common characteristic of producing a protein of this invention.

The genes of this invention may be employed, as will be recognized by the skilled artisan, to produce plasmids or other vectors which, in turn are useful for transforming organisms such as E. coli to produce novel strains of this bacteria which will express the proteins of the invention.

Inhibition of the binding of S. aureus to endothelial cells is a major factor in preventing infection. Accordingly antibodies to proteins expressed by clone 36, its segments clone 14 and clone 30 and even smaller segments are important factors in controlling infection. The proteins and protein segments of this invention are therefore useful to form vaccines to inhibit S. aureus infections of mammals, including humans, by administering an amount of the selected protein which will stimulate the production of protective quantities of antibodies

to limit adhesion of S. aureus. The proteins may also be used to produce antibodies in vitro which may be employed for passive immunization.

The proteins and polypeptide or peptide segments of this invention may be obtained by any of a number of known process including the recombinant DNA techniques described above.

Polypeptide and peptides within the scope of the invention containing, for example from about 6 to 20 or more amino acid segments, may be synthesized be standard solid phase procedures with appropriate amino acids using the protection, deprotection and cleavage techniques and reagents appropriate to each specific amino acid or peptide. A combination of manual and automated (e.g., Applied Biosystem 430A) solid phase techniques can be used to synthesize the novel peptides of this invention. Although less convenient, classical methods of peptide synthesis can also be employed. For background on solid phase techniques, reference is made to Andreu, D., Merrifield, R.B., Steiner, H. and Boman, H.G., (1983) Proc. Natl. Acad. Sci USA 80, 6475-6479; Andreu, D. Merrifield, R.B., Steiner, H. and Boman, H.G., (1985) Biochemistry 24, 1683-1688; Fink, J., Boman, A., Boman, H.G., and Merrifield, R.B., (June 1989) Int. J. Peptide Protein Res. 33, 412-421; Fink, J. Merrifield, R.B., Boman, A. and Boman, H.G., (1989) J. Biol. Chem. 264-6260-6267; each of which being hereby incorporated herein by reference.

The products of the invention are amphoteric. They can exist and be utilized as free bases or as pharmaceutically acceptable metallic or acid addition salts. Suitable metallic salts include alkali and alkaline earth metal salts, preferably

sodium or potassium salts. Acid addition salts may be prepared from a wide variety of organic and inorganic acids including mineral acids, for example citric, lactic, maleic, tartaric, phosphoric and hydrochloric acids. These salts can be prepared by procedures well known to those skilled in the art.

For use as a vaccine, it is presently preferred to administer the selected product in a pharmaceutically acceptable carrier such as a buffer. Mice or other mammals, including humans, when so immunized are protected from colonization and subsequent infection by S. aureus.

Typically, the patient to be protected will be treated with product of the invention in an amount which is effective to elicit a protective immune response. The selected agent may be administered alone or in a pharmaceutically acceptable liquid or solid carrier in which it may be dispersed, dissolved or suspended. If, for example, the patient is to be treated intravenously, the peptide may be suspended as a free base or dissolved as a metallic salt in isotonic aqueous buffer. Other methods of treatment and pharmaceutically acceptable carriers will be apparent to the skilled artisan.

The proteins, polypeptides and peptides of this invention and the genes or oligonucleotides which are employed in their expression are useful as probes for genes and proteins. They are also useful to raise antibodies by which specific strains of streptococci can be identified. The sequence of nucleotides which elicit the unique segment from position 409 to position 419 and modification of this sequence are especially useful as diagnostic probes to identify S. aureus strains. The procedure

is well known to the skilled artisan for identifying other infectious organisms. It involves the preparation of labeled oligonucleotides which aRe used to probe the DNA released from the suspected gram positive bacteria by all lysis, either mechanically or enzymatically.

The following citations are mentioned in the application. They are incorporated by reference.

CITATIONS

1. Blake, M.S., K.H.Johnston, G.J.Russell-Jones, and E.C. Gotschlich. 1984. A rapid sensitive method for detection of alkaline phosphatase-conjugated anti-body on Western blots. Anal.Biochem 136:175-179.

2. Boden, M.K. and J.I. Flock. 1989. Fibrinogen-binding protein/clumping factor from Staphylococcus aureus. Infect.Immun. 57:2358-2363.

3. Cheung, A.L. and V.A. Fischetti. 1989. The role of surface proteins in staphylococcal adherence to fibers in vitro. J.Clin.Invest. 83:2041-2049.

4. Cheung, A.L. and V.A. Fischetti. 1990. The role of fibrinogen in staphylococcal adherence to catheters in vitro. J.Infect.Dis. 161:1177-1186.

5. Cheung, A.L., J.M. Koomey, C.A. Butler, S.J. Projan, and V.A. Fischetti. 1992. Regulation of exprotein expression in Staphylococcus aureus by a locus (sar) distinct from agr. Proc.Natl.Acad.Sci.USA.89:6462-6466.

6. Cheung, A.L., J.M. Koomey, S.Lee, E.A. Jafe, and V.A. Fischetti. 1991. Recombinant human tumor necrosis factor alpha promotes adherence of Staphylococcus aureus to endothelial cells. Infect.Immun.59:3827-3831.

7. Cheung, A.L., M.Krishman, E.A. Jaffe, and V.A. Fischetti. 1991. Fibrinogen acts as a bridging molecule in the adherence of Staphylococcus aureus to culture human endothelial cells. J. Clin.Invest. 87:2236-2245.

8. Cohen, M.L. 1992. Epidemiology of drug resistance: implications for a post-antibiotic era. Science 257:1050-1055.

9. Devereux, J.,P. Haeberli, and O. Smithies. 1984. A comprehensive set of sequences of analysis programs for the VAX. Nucl.Acids Res. 12:387-395.

10. Easom, C.S.F. and C. Adlam. 1983. Staphylococci and staphylococcal infections, Academic Press, New York.

11. Espersen,F.,I. Clemmensen, and V. Barkholt. 1985. Isolation of Staphylococcus aureus clumping factor. Infect.Immun. 49:700-708.

12. Feinberg, A.P. and B. Vogelstein. 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal.Biochem. 132:6-13.

13. Fischetti, V.A., V. Pancholi, and O. Schneewind. 1990. Conservation of a hexapeptide sequence in the anchor region of surface proteins from gram-bacteria. Mol.Microbiol. 4:1603-1605.

14. Fischetti,V.A., V.Pancholi, P.Sellers, J.Schmidt, G.Landau, X.Xu, and O. Schneewind. 1992.Streptococcal M protein: A common structural motif used by gram positive bacteria for biologically active surface molucules, p.31-38. In T.K. Korhonen, T.Hovi, and P.H. Malela (ed.), Molecular recognition in host-parasite interactions, Plenum Press, New York.

15. Garnier, J.,D.J. Osguthorpe, and B.Robson. 1978. Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins.J.Mol.Biol.120:97-120.

16. Halvorson,H.O. and D. Perlman. 1983. A putative signal peptidase recognition site and sequence in eukaryotic and prokaryotic signal peptides.J.Mol.Biol. 167:391-409.

17. Hemker, H.C., M.Bas, and A.D.Muller. 1975. Biochim Biophys.Acta 379:180-188.

18. Kaida, S.T.Miyata, Y.Yoshizawa, H.Igarashi, and S.Iwanaga. 1989. Nucleotide and deduced amino acid sequences of staphylocoagulase gene from Staphylococcus aureus strain 213. Nucl.Acids Res.17:8871.

19. Kaida,S., T.Miyata, Y.Yoshizawa, S.Kawabata, T.Morita, H.Igarashi, and S. Iwanaga. 1987. Nucleotide sequence of the staphylocoagulase gene: its unique COOH-terminal 8 tandem repeats. J.Biochem.102:1177-1186.

20. Kyte, J. and R.F. Doolittle. 1982. A simple method for displaying the hydropathic character of a protein. J.Mol.Biol.157:105-132.

21. Laemmli, U.K. 1970. Cleavage of structure proteins during the assembly of the head of bacteriophage T4, Nature 227:680-685.

22. Lupas, A.,M. Van Dyke, and J. Stock. 1991. Predicting coiled coils from protein sequences. Science 252:1162-1164.

23. Manaitis, T.,E.F. Fritsch, and J. Sambrook. 1989. Molecular cloning, a laboratory manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

24. Neu,H.C. 1992. The crisis in antibiotic resistance. Science 257:1064-1072.

25. Pearson, W.R. and D.L. Lipman. 1988. improved tolls for biological sequence comparison. Proc.Natl.Acad.Sci.USA.88:4781-4785.

26. Phonimdaeng,P.,M. O'Reilly, P. Nowlan, A.J. Bramley, and T. Foster. 1990. The coagualse of Staphylococcus aureus 8325-4. Sequence analysis and virulence of site-specific coagualse-deficient mutants.Mol.Microbiol. 4:393-404.

27. Sanger, F.,S. Nicklen, and A.R. Coulson. 1977. DNA sequencing with chain terminating inhibitors.
Proc.Natl.Acad.Sci.USA.74:5463.

28. Schneewind, O.,P.Model, and V.A. Fischetti. 1992. Sorting of protein A to the staphylococcal cell wall. Cell 70:267-281.

29. So, M., J.H. Crosa, and S. Falkow. 1975. Polynucleotide sequence relationships among Ent plasmids and the relationship between Ent and other plasmids. J. Bacteriol. 121:234-238.

30. Towbin, H., T. Staehelin, and J. Gordon. 1970. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets:Procedure and some applications. Proc.Natl.Acad.Sci.USA. 76:4350-4354.

31. Usui, Y. 1986. Biochemical properties of fibrinogen binding (clumping factor) of the staphylococcal cell surface. Zentralbl.Bakteriol.Mikrobiol.Hyg.[A]262:287-297.

32. Voller, A.,D.E. Bidwell, and A. Barlett. 1976. Enzyme immunoassays in diagnostic medicine. Bull.World Health Organ. 53:55-65.

33. Waldvogel, F.A. 1985. Staphylococcus aureus, p.1097-1116. In G.L. Mandell, R.G.Jr. Douglas, and J.E. Bennett (ed.), Principles and Practice of Infectious Diseases, John Wiley & Sons, New York.

34. Watanakunakorn, C. and I.M. Baird. 1977. Staphylococcus aureus bactermia and endocarditis associated with a removable infected intravenous device. Am.J.Med. 63:253-256.

35. Yother, J. and D.E. Briles. 1992. Structual properties and evolutionary relationship of PspA, a surface protein of Streptococcus pneumoniae, as revealed by sequence analysis J.Bacteriol. 174:601-609.

WHAT IS CLAIMED IS:

1. A gene of the group clone 14, clone 30 and clone 36 which expresses a fibrinogen binding surface protein on S. aureus and segments and analogs of said gene capable of expressing surface proteins on S. aureus having substantially the same activity.

2. A surface protein expressed by a gene of claim 1.

3. A plasmid vector carrying a gene of claim 1.

4. A microorganism transformed by a gene of claim 1 and capable of expressing a fibrinogen binding protein.

5. An antibody to a surface protein of claim 2.

6. A vaccine effective to inhibit the adhesion of S. aureus to fibrinogen because it contains a protein of claim 2.

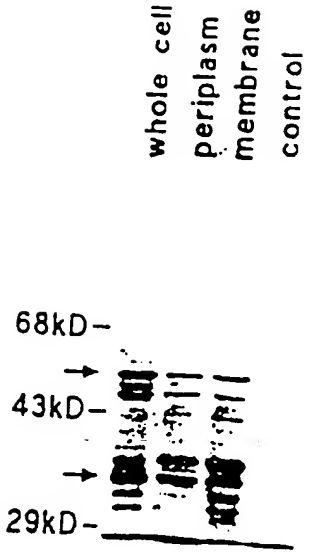
whole cell
periplasm
membrane
control
68kD–
43kD–
29kD–

Fig. 1

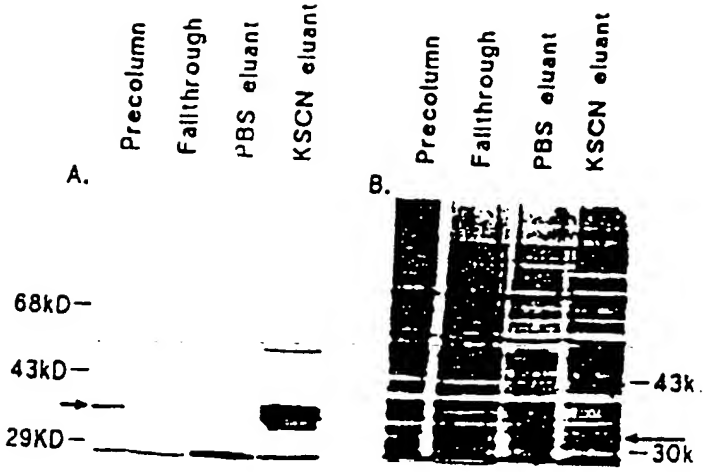
A.
Precolumn
Fallthrough
PBS eluant
KSCN eluant
68kD–
43kD–
29KD–
B.
Precolumn
Fallthrough
PBS eluant
KSCN eluant
–43k
–30k

Fig. 2

```
          10         20         30         40         50         60
          |          |          |          |          |          |
ATGAAAAAGCAAATAATATTCCTAGGCGCATTAGCAGTTGCATCTAGCTTATTTACATGG
METLysLysGlnIleIlePheLeuGlyAlaLeuAlaValAlaSerSerLeuPheThrTrp

          70         80         90         100        110        120
          |          |          |          |          |          |
GATAACAAAGCAGATGCGATAGTAACAAAGGATTATAGTGGGAAATCACAAGTTCAAAAA
AspAsnLysAlaAspAlaIleValThrLysAspTyrSerGlyLysSerGlnValGlnLys

          130        140        150        160        170        180
          |          |          |          |          |          |
AAAGAGCGCCAAAATAGTGCTTCAATATCAGATAGTTATTATTGGGATATCATTAAAAAT
LysGluArgGlnAsnSerAlaSerIleSerAspSerTyrTyrTrpAspIleIleLysAsn

          190        200        210        220        230        240
          |          |          |          |          |          |
CTAGAATTACAATTTACTGCTGCATTAGATTTATTAGAAGATTATCGATATGGTGAAAAA
LeuGluLeuGlnPheThrAlaAlaLeuAspLeuLeuGluAspTyrArgTyrGlyGluLys

          250        260        270        280        290        300
          |          |          |          |          |          |
GAGTATGAAAAAGCAAAAGATCAACTAATGACAAGGATATTAAGTGAAGTCAAGTATTTA
GluTyrGluLysAlaLysAspGlnLeuMETThrArgIleLeuSerGluValLysTyrLeu

          310        320        330        340        350        360
          |          |          |          |          |          |
CTTGAGCAAAAAATTAAAGAATATGACAAGTATAAAGATTTATATAAAGAATATATGAGT
LeuGluGlnLysIleLysGluTyrAspLysTyrLysAspLeuTyrLysGluTyrMETSer

          370        380        390        400        410        420
          |          |          |          |          |          |
AAAAATCCAACGTCAAAGGTAAAAAGAGCAAATTTTGATCAATATAATATCGAAGACCTA
LysAsnProThrSerLysValLysArgAlaAsnPheAspGlnTyrAsnIleGluAspLeu

          430        440        450        460        470        480
          |          |          |          |          |          |
AGAGAAAAAGAATATAATGATTTACTAAGTTCTATTAAAGATGCGGTAGAAACATTTAAA
ArgGluLysGluTyrAsnAspLeuLeuSerSerIleLysAspAlaValGluThrPheLys

          490        500        510        520        530        540
          |          |          |          |          |          |
TCAGATGTACAAAAAATAGAATATGAAAATAAAGAGTTAAAATCTTATTCTTACGAAGAA
SerAspValGlnLysIleGluTyrGluAsnLysGluLeuLysSerTyrSerTyrGluGlu

          550        560        570        580        590        600
          |          |          |          |          |          |
GAAAAGAAGGCTGCTTCTAGAGTTGATGATTTAGCAAATAAAGCGTATAGCGTTTATTTT
GluLysLysAlaAlaSerArgValAspAspLeuAlaAsnLysAlaTyrSerValTyrPhe

          610        620        630        640        650        660
          |          |          |          |          |          |
GCGTTTGTTAGGGATACACAACATAAAACTGAGGCATTAGAGTTAAAAGCGAAAGTGGAT
AlaPheValArgAspThrGlnHisLysThrGluAlaLeuGluLeuLysAlaLysValAsp
```

```
          670         680         690         700         710         720
          |           |           |           |           |           |
TTAGTTTTAGGTGATGAGGACAAACCACATCGCATCTCTAATAAAAGAATTGAAAATGAA
LeuValLeuGlyAspGluAspLysProHisArgIleSerAsnLysArgIleGluAsnGlu

          730         740         750         760         770         780
          |           |           |           |           |           |
ATGCTAAAAGATTTAGAATCTATTATTGAAGATTTCTTTATAGAAACTGGTTTAAATAAA
METLeuLysAspLeuGluSerIleIleGluAspPhePheIleGluThrGlyLeuAsnLys

          790         800         810         820         830         840
          |           |           |           |           |           |
CCTGATAATATTACGAGTTATGATAGTAGTAAACATGACTATAAACATCATAGAGAAGGC
ProAspAsnIleThrSerTyrAspSerSerLysHisAspTyrLysHisHisArgGluGly

          850         860         870         880         890         900
          |           |           |           |           |           |
TTTGAGGCTCTAGTTAAAGAAACAAGAGAAGCCGTCGATAAGGCTGACGAATCTTGGAAA
PheGluAlaLeuValLysGluThrArgGluAlaValAspLysAlaAspGluSerTrpLys

          910         920         930         940         950         960
          |           |           |           |           |           |
ACTAAAACTGTCAAAACATATGGTGAAGCTGAAACAAAAGCACATGTTGTAAAAGAAGAG
ThrLysThrValLysThrTyrGlyGluAlaGluThrLysAlaHisValValLysGluGlu

          970         980         990         1000        1010        1020
          |           |           |           |           |           |
AAGAAAGTTGAAGAACCTCAAGCACCTAAAGTTGGAAACCAGCAAGGGGATAAAACTACA
LysLysValGluGluProGlnAlaProLysValGlyAsnGlnGlnGlyAspLysThrThr

          1030        1040        1050        1060        1070        1080
          |           |           |           |           |           |
GTTGATAAAGAAGTTGAAGCAACACAACCAGTGGCACAGCATTTAGTTGGAATTCCACAG
ValAspLysGluValGluAlaThrGlnProValAlaGlnHisLeuValGlyIleProGln

          1090        1100        1110        1120        1130        1140
          |           |           |           |           |           |
GGCACAATTACAGGTGAAATTGTAAAAGGTCCAGACTATCCAACGATGGAAAATAAAATG
GlyThrIleThrGlyGluIleValLysGlyProAspTyrProThrMETGluAsnLysMET

          1150        1160        1170        1180        1190        1200
          |           |           |           |           |           |
TTACAAGGTGAAATAGTTCAAGGTCCAGATTTTCTAACAATGGAACAAAACAGACCATCT
LeuGlnGlyGluIleValGlnGlyProAspPheLeuThrMETGluGlnAsnArgProSer

          1210        1220        1230        1240        1250        1260
          |           |           |           |           |           |
TTAAGTGACAATTATACACAACCATCTGTGACTTTACCGTCAATTACAGGTGAAAGTACA
LeuSerAspAsnTyrThrGlnProSerValThrLeuProSerIleThrGlyGluSerThr

          1270        1280        1290        1300        1310        1320
          |           |           |           |           |           |
CCAACGAACCCTATTTTAGAAGGTCTTGAAGGTAGCTCATCTAAACTTGAAATAAAACCA
ProThrAsnProIleLeuGluGlyLeuGluGlySerSerSerLysLeuGluIl LysPro
```

```
        1330        1340        1350        1360        1370        1380
         |           |           |           |           |           |
CAAGGTACTGAATCAACGTTGAAAGGTACTCAAGGAGAATCAAGTGATATTGAAGTTAAA
GlnGlyThrGluSerThrLeuLysGlyThrGlnGlyGluSerSerAspIleGluValLys

        1390        1400        1410        1420        1430        1440
         |           |           |           |           |           |
CCTCAAGCAACTGAAACAACAGAAGCTTCTCAATATGGTCCGAGACCGCAATTTAACAAA
ProGlnAlaThrGluThrThrGluAlaSerGlnTyrGlyProArgProGlnPheAsnLys

        1450        1460        1470        1480        1490        1500
         |           |           |           |           |           |
ACACCAAAGTATGTGAAATATAGAGATGCTGGTACAGGTATCCGTGAATACAACGATGGA
ThrProLysTyrValLysTyrArgAspAlaGlyThrGlyIleArgGluTyrAsnAspGly

        1510        1520        1530        1540        1550        1560
         |           |           |           |           |           |
ACATTTGGATATGAAGCGAGACCAAGATTCAACAAGCCAAGTGAAACAAATGCATACAAC
ThrPheGlyTyrGluAlaArgProArgPheAsnLysProSerGluThrAsnAlaTyrAsn

        1570        1580        1590        1600        1610        1620
         |           |           |           |           |           |
GTAACGACAAATCAAGATGGCACAGTAACATATGGCGCTCGCCCGACACAAAAAAAGCCA
ValThrThrAsnGlnAspGlyThrValThrTyrGlyAlaArgProThrGlnLysLysPro

        1630        1640        1650        1660        1670        1680
         |           |           |           |           |           |
AGCAAAACAAATGCATACAACGTAACAACACATGCAAATGGTCAGATATCATATGGCGCT
SerLysThrAsnAlaTyrAsnValThrThrHisAlaAsnGlyGlnIleSerTyrGlyAla

        1690        1700        1710        1720        1730        1740
         |           |           |           |           |           |
CGCCCGACACAAAAAAAGCCAAGCAAAACAAATGCATACAACGTAACAACACATGCAAAT
ArgProThrGlnLysLysProSerLysThrAsnAlaTyrAsnValThrThrHisAlaAsn

        1750        1760        1770        1780        1790        1800
         |           |           |           |           |           |
GGTCAAGTATCATATGGCGCTCGCCCGACACAAAAAAAGCCAAGCAAAACAAATGCATAT
GlyGlnValSerTyrGlyAlaArgProThrGlnLysLysProSerLysThrAsnAlaTyr

        1810        1820        1830        1840        1850        1860
         |           |           |           |           |           |
AACGTAACAACACATGCAAATGGTCAAGTATCATACGGAGCTCGCCCGACATACAAGAAG
AsnValThrThrHisAlaAsnGlyGlnValSerTyrGlyAlaArgProThrTyrLysLys

        1870        1880        1890        1900        1910        1920
         |           |           |           |           |           |
TCAAGCGAAACAAACGCATATAACGTAACAACACATGCAGATGGTACTGCGACATATGGG
SerSerGluThrAsnAlaTyrAsnValThrThrHisAlaAspGlyThrAlaThrTyrGly

        1930
         |
CCTAGAGTAACAAAATAA
ProArgValThrLys***
```

```
      a   b   c   d   e   f   g
  1                                  Met Lys Lys Gln Ile Ile Phe Leu Gly Ala
 11                                  Leu Ala Val Ala Ser Ser Leu Phe Thr Trp
 21                                  Asp Asn Lys Ala Asp Ala Ile Val Thr Lys
 31                                  Asp Tyr Ser Gly Lys Ser Gln Val Gln Lys
 41                                  Lys Glu Arg Gln Asn Ser Ala Ser Ile Ser
 51                                  Asp Ser Tyr Tyr Trp Asp Ile
 58  Ile Lys Asn Leu Glu Leu Gln
 65  Phe Thr Ala Ala Leu Asp Leu
 72  Leu Glu Asp Tyr Arg Tyr Gly
 79  Glu Lys Glu Tyr Glu Lys Ala
 86  Lys Asp Gln Leu
 90  Met Thr Arg Ile Leu Ser Glu
 97  Val Lys Tyr Leu
101              Leu Glu Gln Lys
105  Ile Lys Glu Tyr Asp Lys
111  Tyr Lys Asp Leu
115  Tyr Lys Glu Tyr Met Ser Lys
122          Asn Pro Thr Ser Lys
127  Val Lys Arg Ala Asn
132  Phe Asp Gln Tyr Asn
137  Ile Glu Asp Leu Arg Glu Lys
144          Glu Tyr Asn Asp Leu
149  Leu Ser Ser Ile Lys Asp Ala
156  Val Glu Thr Phe Lys Ser Asp
163  Val Gln Lys Ile Glu Tyr Glu
170  Asn Lys Glu Leu Lys Ser Tyr
177          Ser Tyr Glu Glu Glu
182  Lys Lys Ala Ala Ser Arg
188  Val Asp Asp Leu Ala Asn Lys
195  Ala
196
206                              Tyr Ser Val Tyr Phe Ala Phe Val Arg Asp
216                              Thr Gln His Lys Thr Glu Ala Leu Glu Leu
226                              Lys Ala Lys Val Asp Leu Val Leu Gly Asp
234                              Glu Asp Lys Pro His Arg Ile Ser
                     Asn Lys Arg
237  Ile Glu Asn Glu Met
242  Leu Lys Asp Leu Glu Ser Ile
249  Ile Glu Asp
252
262                             Phe Phe Ile Glu Thr Gly Leu Asn Lys Pro
                          Asp Asn
264  Ile Thr Ser Tyr Asp Ser Ser
271  Lys His Asp Tyr Lys His His
278  Arg Glu Gly Phe Glu Ala Leu
285  Val Lys Glu Thr Arg Glu Ala
292  Val Asp Lys Ala Asp Glu
```

Fig. 4

```
                451
DB              LKGtQGESSD IEVKPQATET TEASQYGPRP QFNKTPKYVK YRDAGTGIRE      500
8325-4          LKGtQGESSD IEVKPQATET TEASQYGPRP QFNKTPKYVK YRDAGTGIRE
213             LKGiQGESSD IEVKPQATET TEASQYGPRP QFNKTPKYVK YRDAGTGIRE
BB              LKGiQGESSD IEVKPQATET TEASQYGPRP QFNKTPKYVK YRDAGTGIRE

                501
DB              YNDGTFGYEA RPRFNKPSET NAYNVTTNQD GTVtYGARPT Q.........      550
8325-4          YNDGTFGYEA RPRFNKPSET NAYNVTThan GqVSYGARPT ykKPSETNAY
213             YNDGTFGYEA RPRFNKPSET NAYNVTTNQD GTVSYGARPT QNKPSETNAY
BB              YNDGTFGYEA RPRFNKPSET NAYNVTTNQD GTVSYGARPT QNKaSETNAY

                551
DB              .......... .......... .......... .......... ..........      600
8325-4          NVTTHANGQV SYGARPTQ.. .......... .......... ..........
213             NVTTHANGQV SYGARPTQ.. .......... .......... ..........
BB              NVTTHANGQV SYGARPTQkk psetnaynvt thangqvsyg arptynkpse

                601
DB              .......... .......... ..kKPSKTNA YNVTTHANGQ ISYGARPTQk      650
8325-4          .......... .......... ..nKPSKTNA YNVTTHgNGQ VSYGARqaQn
213             .......... .......... ..nKPSKTNA YNVTTHANGQ VSYGARPTQk
BB              tnaynvtthg ngqvsygarp tykKPSKTNA YNVTTHANGQ VSYGARPTQn

                651
DB              KPSKTNAYNV TTHANGQVSY GARPTqKKPS kTNAYNVTTH ANGQVSYGAR      700
8325-4          KPSKTNAYNV TTHANGQVSY GARPTy.... .......... ..........
213             KPSKTNAYNV TTHANGQVSY GARPTyKKPS ETNAYNVTTH ANGQVSYGAR
BB              KPSeTNAYNV TTHANGQVSY GARPTqnKPS ETNAYNVTTH gNGQVSYGAR

                701                 730
DB              PTYKKsSeTN AYNVTTHADG TATYGPRVTK
8325-4          ...KKPSkTN AYNVTTHADG TATYGPRVTK
213             PTqKKPSeTN AYNVTTHADG TATYGPRVTK
BB              PTYnKPSkTN AYNVTTHADG TATYGPRVTK
```

Fig. 5

Strains

```
                1
DB              MKKQIIfLGA LAVASSLFTW DNKADAIVTK DYSgkSqVqk KerqNsasIS 50
8325-4          MKKQIISLGA LAVASSLFTW DNKADAIVTK DYSgkSqVNa gSK.NGtLId
213             MKKQIISLGA LAVASSLFTW DNKADAIVTK DYSkeSrVNE KSK.kGatVS
BB              MKKQIISLGA LAVASSLFTW DNKADAIVTK DYSkeSrVNE nSK.yGtLIS

               51
DB              DsYyWdIikn LElQFtaAlD LLEdYrYGEK eYekAKDqLM TRILsEvkYL 100
8325-4          sRYlnsalyy LEdyiiyAig LtnkYeYGDn iYKEAKDRLL ekVLrEDQYL
213             DYYyWkIidS LEaQFtgAiD LLEnYkYGDp iYKEAKDRLM TRVLGEDQYL
BB              DWYlkgrltS LEsQFinAlD iLEtYhYGEK eYKDAKDkLM TRILGEDQYL

              101
DB              LEqKikeYDk YKdlYKeYms kNPTSK.vKr AnFdqYNiED LreKEYNDLL 150
8325-4          LERKksqYED YKqwYanYKK ENP.rtdLKM AnFhkYNlEE LsMKEYNELq
213             LkkKideYEl YKKwYKssnK nt......nM ltFhkYNlyn LTMnEYNDiF
BB              LERKkvqYEE YKKlYqkYKe ENPTSKgLKL ktFdqYtiED LTMrEYNELt

              151
DB              sSiKDAVEtF ksDVqkIEye NkELKsYsyE EEkKAasrVd DLankaysvy 200
8325-4          DaLKrAlDDF hrEVKdIkdK NsDLKtFnaa EEDKATkeVy DLvseidtlv
213             nSLKDAVyqF nKEVKeIEhK NvDLKqFdkD gEDKATkeVy DLvseidtlv
BB              ESLKsAVkDF eKDVekIEnq hhDLKpFtdE mEEKATsrVd DLankaysvy

              201
DB              fAFvRDtqhk teAlELkAKv DLVLGDeDkP HrIsNkRIen EMlkDLeSII 250
8325-4          vsYygDkdyg ehAkELrAKl DLILGDtDnP HkItNERIkK EMIdDLnSII
213             vtYyaDkdyg ehAkELrAKl DLILGDtDnP HkItNERIkK EMIdDLnSII
BB              fAFvRDtqhk teAlELkAKv DLVLGDeDkP HrIsNERIeK EMIkDLeSII

              251
DB              EDFFiETglN kPdnITsYDs sKHdYKhHrE ...gFEaLVk ETreAVdkAD 300
8325-4          DDFFmETkqN rPksITkYnp ttHNYKtnSD NKPnFDkLVe ETkkAVKEAD
213             DDFFmETkqN rPnsITkYDp tKHNFKekSE NKPnFDkLVe ETkkAVKEAD
BB              EDFFiETglN kPgnITsYDs sKHhYKnHSE ...gFEaLVk ETreAVanAD

              301
DB              ESWKTKTVKt YGEaETKahV VKEEKKVEEP QAPKVgNQQg dKTTvdKevE 350
8325-4          DSWKkKTVKK YGETETKSPV VKEEKKVEEP QAPKVdNQQE VKTTAGKAEE
213             ESWKnKTVKK YeETvTKSPV VKEEKKVEEP QlPKVgNQQE VKTTAGKAEE
BB              ESWKTKTVKK YGEsETKSPV VKEEnKVEDP QsPKfdNQQE VKTTAGKAEE

              351
DB              aTQPVAQhLV gIPQGTITGE IVKGPDYPTM ENKmLQGEIV QGPDFLTMEQ 400
8325-4          TTQPVAQPLV KIPQGTITGE IVKGPEYPTM ENKTvQGEIV QGPDFLTMEQ
213             TTQPVAQPLV KIPQeTIyGE tVKGPEYPTM ENKTLQGEIV QGPDFLTMEQ
BB              TTQPVAQPLV KIPQGTITGE IVKGPEYPTM ENKTLQGEIV QGPDFpTMEQ

                      ↓          ↓
              401     ____________
DB              nrPSLSDNYT QPsvtlpsit gesTptNPIL EGLEGSSSKL EIKPQGTEST 450
8325-4          sgPSLSnNYT nP..... ...plTNPIL EGLEGSSSKL EIKPQGTEST
213             nrPSLSDNYT QP..... ...TTpNPIL EGLEGSSSKL EIKPQGTEST
BB              sgPSLSDNYT QP..... ...TTpNPIL EGLEGSSSKL EIKPQGTEST
```